Where are we now in British health economics?

Mark Blaug

¿En qué estamos en la economía de salud británica?

Mark Blaug
Health economics took off in 1970 or thereabouts, just after the take-off date for the economics of education. Although early health economics made use of human capital theory as did the economics of education, it soon took a different route inspired by Arrow’s work on medical insurance. The economics of education failed to live up to its promising start in the 1960s and gradually ran out of steam. The economies of health, however, has made steady theoretical and empirical progress since 1970 principally in coming to grips with the implications of supplier-induced demand and the difficulties of evaluating health care outcomes. Some of the best work on British health economics has been in the area of normative welfare economies, defining more precisely what is meant by equity in the delivery of health care and measuring the degree of success in achieving equity. Recent efforts to reform the NHS by the introduction of “quasi markets” have improved the quality and quantity of health care in Britain. In short, British health economics has been characterized by the use of Pigovian piecemeal rather than Paretian global welfare economics, retaining a distinctive style that sets it apart from American health economics. ©1998 John Wiley & Sons, Ltd.
It is my task to review the development of British health economics since about 1970, which is not only the date of the founding of the UK Health Economists’ Study Group but, roughly speaking, the take-off date for health economics in the UK [3]. My review is limited to developments in the UK and that presents something of a problem because the centre of gravity of the subjects of health economics is, surely, somewhere between the Eastern and Western seaboards of the United States. The explanation for the US dominance of health economics is, I believe, two-fold, in part it is because American economists dominate the whole of economics, overwhelming British economics by sheer quantity of not quality - how can we compete with a system that produces over 800 PhDs in economics year in and year out [4]? Furthermore, the fact that medical care in America is not only privately provided but largely privately financed via private medical insurance allows the applications of such standard economic concepts as market-clearing prices and utility-maximising economic agents to problems of the health care. By way of contrast, the reliance of the NHS on public provision and public finance does not provide an ideal terrain for economics analysis. Nevertheless, the economics of health in Britain, while it continues to be overshadowed by American health economics, has not been overpowered by it and to this day retains its own distinctive, voice and style. In short, it is possible to look at health economics in Britain as a separate sub-discipline and to consider its progress or lack of it since 1970.

**Is there a central paradigm?**

Victor Fuchs began its excellent encyclopaedia entry on “Health Economics” in *The New Palgrave* [5] with the pronouncement:

“Health economics is an applied field in which empirical research dominates. It draws its theoretical inspiration principally from four traditional areas of economics: finance and insurance, industrial organization, labour and public finance. Some of the most useful work employs only elementary economic concepts, but requires detailed knowledge of health technology and institutions. Policy-oriented research plays a major role and many important policy-relevant articles are published in journals read by physicians and others with direct involvement in health”.

One can see exactly what Fuchs is driving at and yet it seems, to go too far in denying a dominant intellectual paradigm in health economics. To be sure, there was nothing in health economics quite like the human-capital concept in the economies of education that sparked off what has been called “the human investment revolution in economic thought” in 1960 or thereabouts. But perhaps Arrow’s seminal 1963 paper inspired budding health economists almost as much as did papers by Jacob Mincer and Gary Becker in the famous Supplement volume of the *Journal of Political Economy* in October 1962.

If education was to be viewed henceforth as an investment in the extra future earnings of more educated people, so was health to be viewed henceforth as a special kind of commodity whose future maintenance at an uncertain date and at an uncertain cost necessarily required either public provision free at the point of use or medical insurance privately or publicly financed, and with insurance came the twin evils of “moral hazard” and “adverse selection”, market failures that could be mitigated by such devices as deductibles and coinsurance but which could never be entirely cured. In other words, Fuch’s words notwithstanding, health
economics is more than even physical investment itself, beset with uncertainty and the problems of dealing with uncertainty.

In addition to the presence of pervasive uncertainty, but associated with it is the presence of a remarkable degree of ignorance the part of consumers of health care about both the quantity of health care the require and the quality of the care they receive. Now, of the course, consumer ignorance is a major source of market failure in the provision of many goods and services, with welfare economists usually note briefly only to pass on to more exciting aspects of market failure. However, the field of health care is characterised by endemic and irre-mediable consumer ignorance and is indeed the best illustration, in economics of the phenomena of asymmetric information and missing markets. Paretian welfare economics is grounded on a number of premises of which the most important is the notion of consumer sovereignty, namely, that every individual is the best judge and indeed the only judge of their own welfare: is follows that “willingness-to play’ is the measure of the marginal benefit of a commodity to individuals and hence that social welfare is simply the sum of the explicit or implicit willingness-to-pay of individuals [6]. This master-premise of Paretian welfare economics is more clearly violated in the case of health care than in the provision of any other commodity or service that comes to mind, including that of the education. It would seem therefore that Paretian welfare economics is either irrelevant or at least of dubious value in assessing the efficiency of health care provision [7].

Be that as is may, it is hardly surprising that this is a conclusion that most health economists have been loath to swallow. Arow’s trail-blazing 1963 paper underlined the alarming ignorance of patients which forces them to delegate all vital decisions in medical treatment, to the doctor that is treating them. It is noteworthy that this realistic analysis of the principal agent relationship in health care nevertheless followed Arrow’s standard exposition of the Pareto optimal properties of efficient allocation of resources in health care [1]. As a consequence, some early readers of Arrow failed to notice that the thrust of his essay was to show that health care markets invariably fail and that the best we can do is to minimize the consequences of market failure in health by various norms and conventions: what we can never do is to eradicate entirely the inherent inefficiencies of resources allocation in health.

Similarly, there is widespread recognition among health economists of the phenomenon of “supplier induced demand”, that is, the proposition that fee-for services paid doctors can induce demand for their services, and that even salaried doctors in the NHS are inclined, at least to some extent, to do the same. But if the demand of health care is inseparable from the supply or health care, the economic analysis of health care will have to reconsider many of the tools that economists traditionally employ. Recognition of this fact no doubt explains why Michael’s Grossman’s application of human capital theory to health [8] attracted few adherents and has been more often attacked than endorsed [9-13]. Grossman models the demand for health care as a perfectly rational decision of incur present health insurance premiums to compensate for the loss of future earnings that would result form illness. Unfortunately, this loss is not risky in the sense that its probability can be calculated; it is radically uncertain in the Knightian, sense and neither the date of the illness nor the cost of the prescribed treatment, not no mention the effectiveness of the treatment
in terms of recovered health, can be predicted even probabilistically.

To sum up, health economics would seem to be a perfect topic for heterodox dissent and yet, surprisingly enough, radical economists and Marxists have not on the whole been attracted to health economics. Still, and this is my main point, health economics is a field which must make the average neoclassical economics squirm because it challenges, his or her standard assumptions at every turn. Perhaps that is precisely what makes it so interesting to study.

**Has there been progress?**

I come now to the main question posed by this survey: has there been progress in health economics? The question of progress in economics is a famous question, which is never easy to answer unambiguously. Progress in economics (or in any other subject for that matter) can take the form either of “Theoretical progress” or “Empirical progress”, or both [14]. By “Theoretical Progress “ I mean greater precision in the definition of terms and the relationship between terms and, in general, improved conceptual clarity, frequently accompanied by analytic innovations: in short, sharper tools for the “box of tools” that is the economic theory according to the young Joan Robinson. Theoretical progress may of may not be accompanied by “Empirical progress”. Which is a much more elusive idea than theoretical progress. By “empirical progress” I mean a firmer grasp, a better explanation, a more accurate prediction of the behaviour of economic factors and the operation of the economic system their actions produce. I know of no economist who denies that the history of economic thought, right up to yesterday, is characterised by theoretical progress but some economists have doubts that there has been significant empiric progress, at least in the intermediate run of the last half century. This is a doubt which I do not share but suffice it to say that such doubts are almost always generated by the somewhat naive belief that the economics is nothing if it, cannot make precise, quantitative predictions of changes in economic events. If instead we set our sights on accurate generic predictions of trends and patterns-qualitative rather than quantitative predictions-then the track-record of empiric progress in economics is nothing to be ashamed of.

With this much taken for grated, let us briefly compare and contrast progress in the economics of education and progress in the economics of health since about 1970. In 1970, I personally believed that the economics of education held out more promise of fruitful research than health economics, even form a political standpoint, it was education and not health that was “sexy” in the 1960s. However, by the mid 1970s, I became increasingly disenchanted by the economics of education, dominated as it was by human capital theory. I have expounded my jaundiced views elsewhere [15] and small say no more about them here. To suggest the flavour of may scepticisms, however, let me simply say that despite considerable empiric work in the economics of education over last three decades, human capital theory has failed in resolve the difficulties in its research agenda that appeared at the very outset, such as the relative impact on individual earnings of endowed ability, acquired ability and education attainment, on the one hand. and quality and quantity of formal schooling and on-the-job training, on the other. The so-called screening or filter hypothesis has never been convincingly tested, and certainly has not been decisively refuted, and screening throws cold water on any belief that the so-
cial rate of return on educational investment can provide governments with an investment criterion for education spending. Nothing new has been said since 1970 on so central a concept as the externalities of education and even production function estimates of education—the relationship between school resources and student achievement—has made little progress since famous Coleman Report of 1966 even from a strictly statistical point of view recent work on education production functions continues to assume the school strive to maximize student achievements scores in cognitive tests, a dubious assumption at best [16, 17]: there are suggestive analogies here with cost studies of hospitals that assume that not-for-profit hospitals strive to minimize the costs of secondary education.

Perhaps the most alarming symptom in this lack of both theoretical and empirical progress in the economics of education over there decades is the fact that virtually all the 100 articles in the 1985 International Encyclopaedia of Education devoted to the economics of education [18], could just as well have been write in 1970 or even 1960 [19]. When we examine the leading American textbook in the field [20] or a recent British textbook [21], the first one in economics of education in Britain for twenty years, one is struck by how little has been added to the subject since the 1960s [22]. Of course there are data in these books refereeing to more recent years but when we ask what we really know about the economic aspect or the consequences of education in 1997 that we did not know Becker created the subject of the economics of education in his great opus Human Capital in 1964, the answer is: very little. A collection of papers on Recent Development in the Economics of Education [23, 24] adds a new emphasis on equity to the older interests in efficiency but, apart from that, it reveals an alarming tendency to shuffle the same set of intellectual cards in more or less the same combinations and permutations. We are driven to the cruel conclusion that the economics of education is now moribund as a subject of study. By way of contrast, I believe that there has been steady theoretical and empirical progress in health economics ever since 1970 as I now hope to show.

Health economics is indeed as Alan Williams has described it, namely, “the cheerful face of the dismal science” [25].

Could the economics of education have benefited from the development of ideas in health economics? Perhaps, but the differences between the two fields have doomed all attempts at drawing them closer together. For one thing, there is not analogy in health economics to the calculation of rates of return on educational investment, which have served economics of education with both an explanation of the private demand for education and an investment criterion for public expending on education [26]. Education for the individual is a kid of insurance against the probability of future unemployment and low earnings and to the extent that education and particularly higher education, is heavily subsidized, moral hazard induces an excess demand for health care. But the rationing of places in upper secondary and higher education have always kept a firm lid on this excess demand and the result is that economic analysis post-compulsory education is more and analysis of the politics than of the economics of education. Even the demand for income-contingent student loans (a real analogy to the demand for private medical insurance) is an unrealised demand, at least at the present juncture, because all British governments. Tory or Labour, have always given greater weight to the political influence of the parents of potentials university students than to the economic arguments in
favour of income-related student loans. In short, the economics of education and the economics of health have gone their separate ways and I see little hope that steady progress in health economics over the last few decades will now at long last breathe fresh life in to the economics of education [27].

**Developments in health economics: supplier-induced demand**

To return to health economics, we begin our history by noting the textbook literature in British health economics because textbooks are true barometers of energy and interest in a subject. The vigorous state of British health economics is revealed by an unusually steady flow of lively textbook, particularly of the type addressed to health care managers and planners on the application of economic appraisal techniques to the health care sector. Early examples are Cullis and West [12] and Mooney, Russell and Weir [28] (see also Evans [29] addressed to Canadian readers). Soon followed by Mooney [30], McGuire, Henderson and Mooney [13], not to mention long chapters in textbooks on social policy issues like Barr [31], Glennerster [32] and, most incisive of all (given the limitation of space), McGuire, Fenn and Mayhew [33]. Finally Mooney and McGuire have edited a whole series of book-length monographs and anthologies of which Donaldson and Gerard [34] is an outstanding example.

Revealing too is the treatment of health economics in textbooks addressed to students of social policy and social administration [35-38] with contrasts with similar treatments of the economics of education that complain of intellectual stagnation in the subject since 1970 [38]. The textbooks literature now reaches down to secondary schools [39] and perhaps nothing better that health economics has arrived as a genuine branch of the economics than its appearance in the A-level Economics syllabus.

We turn now to he substantive issues, beginning with that twin-headed monster of medical insurance moral hazard and adverse selection - which figures so prominently in Arrow’s pioneering article and the discussion which gave rise to it [40, 41]. (Arrow does no actually use the term “adverse selection” but the idea is nevertheless clearly discernible in his essay.) A distinction was soon drawn between ex ante and ex post moral hazard and between “consumer moral hazard” and producer moral hazard” both leading to over-utilization of medical care. Cutting in the opposite direction, adverse selection can lead to “cream skimming” by which insurance companies keep costs down by selecting healthy patients, which may lead to under-utilization of medical care. [42-44]. From there is only a short jump to the concept of supplier-induced demand [45] and the so-called third-party payment problem” [31]. Although user charges in the NHS exist only for dental care, prescription drugs and optician charges, rationing by waiting time and non-provision can cause “demands” for health care in Britain to depend on supply. In short, supplier-induced demand is not entirely banished even in the NHS, but all attempts to quantify it nevertheless suggest that it is a minor problem [13, 30, 46, 47]. There may be consumer hazard but there ought to be little producer moral hazard in tax-financed health care system with salaried doctors [48]. Is the third-party payment problem in fact the root cause of the extraordinary explosion of medical costs throughout the industrialised world in recent decades, and particularly so in the United States where heavy reliance on private medical insurance has produced an staggering 15 % of national income
devoted to medical care? This medical cost containment question has perhaps produced more controversy in health economics than any other issue.

There are at least three strands of opinion on the question. According to Fuchs' survey of 46 leading figures in American health economics, four out of five believed that third-party payment results in patients using services whose costs exceed their benefits, and this excess of costs over benefits amounts to at least 5% of total health care expenditures—a modest version of the thesis [49]. Moreover, two out of three believed in supplier-induced demand, that is, that doctors do have the power to shift the demand curse for their services. Nevertheless, four out of five also believed that the primary reason for the increase in the health sector's share of GDP over the last 30 years was neither the third-party payment problem nor supplier-induced demand but rather the pace and nature of technical progress in medicine. Joseph Newhouse, the American co-editor of the Journal of Health Economics, therefore speaks for most of the profession when he argues [50] that a large part of the demand for medical services in the United States is a genuine demand, reflecting a willingness to pay for ever more costly sophisticated medical technology (see also Weisbrod [51]).

However, even if the quality of medical hardware had stood still, Baumol's cost disease [52], known in Britain as “the relative price effect” [53], would have guaranteed an increase in the real price charged by doctor to patients and in the real price of a hospital bed simply because the rate of growth in productivity in a labour-intensive sector like health care has never matched the productivity growth of other sectors.

Nevertheless, it is only the quantitative impact of the third payment problem and supplier-induced demand that is a bone of contention among American health economists. Everyone agrees that the doctor is a double agent for two principals, that is vis-à-vis the patient vis-à-vis the insurance company [54]. Even with accurate diagnosis of an illness, the relation between health care and health outcomes is so loose that performance guarantees cannot be given to either principal: this is a kind of information asymmetry that faces both ways and that is perhaps even shared by the agent him-or herself. Some American health economists have nevertheless tried to make medical treatment amenable to ordinary economics variables like income and user price by separating the patient's decisions to consult a doctor from the decision to comply with the doctor's advice [55]. But this artificial distinction depends crucially that the potential patient having a threshold health level below which he/she will not consult a doctor. So pervasive is consumer ignorance in this regard, however, that is of doubtful that any such threshold level has sufficient stability to make it a predicable parameter for analysis.

We come in the end, therefore, to the proposition that there is almost certainly supplier-induced demand particularly in a fee-for-service payments system if, for example, doctors seek to maintain a target income and find themselves with fewer patients because of rising fees. Because supplier-induced demand shifts the demand curve for medical care to the right, its effect is identical to the issue of medical insurance or an increase in medical insurance cover.

For that reason, Charles Phelps, a leading American health economist, entitles one of his papers on supplier-induced demand: “Induced demand—can we ever know its extent?” [56] In reaction to this paper, Victor Fuchs [57], notes de eagerness with which many economists seek to pour cold water on
the notion of suppliers-induced demand as subversive of the economic approach to health care [see also 58-61]. But microeconomic behaviour apart, a recent state of the art econometric exercise [62] confirms the view that fee-for-service payments systems as compared to salaried-capitation methods to contribute significantly to the explanation of health care expenditure variation between countries.

We have wandered far from British concerns but it is inevitable that the supplier induced demand issue rouses more furious controversy in the US than in the UK. The NHS and per-capitation payments to doctors do not entirely prevent supplier-induced demand but, nevertheless, they do not induce doctors systematically to encourage demand. However, we do have or rather did have, a fee-for-service payments system for dentists and some observers think they have indeed detected supplier-induced demand in British dental care provision [63, 64]. The notion of supplier-induced demand underlines the central role of asymmetric information and hence the “incompleteness” of the agency relationship in health care markets [65]. The acute nature of the information asymmetry in health matters [31] is frequently played down by advocates of the free market approach to health [66]. Similarly, it is standard practice in textbooks to estimate the welfare burden of excess health insurance in an American-type health care system on the basis of a fully informed consumer demand curve as if consumers could equate the marginal costs and benefit of the health care despite moral hazard induced by insurance [30, 34, 67]. This amounts to denying the existence of supplier-induced demand and ends up minimizing if not actually ignoring the role of asymmetric information in health care markets.

**Evaluating health care outcomes**

We pass from primary care considerations to questions about secondary care, that is, hospital costs for surgical operations and outpatient services. Among the earliest examples of work in health economics was Martin’s Feldstein’s study of production of American and British hospitals [68]. Production function studies of hospital costs have ever since been one of the staples of health economics. I shall little about them here because Wagstaff has surveyed the field up to 1987 with a thoroughness that leaves nothing more to be said. [47]. Nevertheless I must reiterate the well-known objection that the meaning of hospital production functions is problematic, because, firstly, the multidimensional output of hospitals in not in practice closely related to the quantity of health care inputs as measured in production function studies—hence the use of “throughput” measures as a proxy for output in these studies—and, secondly, because even private hospitals, and certainly public hospitals, are rarely profit-maximizing enterprises, as in the case of similar institutions, such, as schools, it is far from obvious what is the maxim and of not-for-profit hospitals. Do they aim to turn away as few patients as possible, minimizing the number of spare beds in the face of a stochastic demand of hospital services, or do they aim to fill all beds all the even if it means turning some patients away? The question is difficult enough to answer in a medical care system with privately insured patients as in the United States [69, 70] but it is even more difficult to answer in our own publicly financed health care system [13]

It is precisely on the question of how to evaluate health care outcomes so as to compare them to their costs in a system that lacks willingness—to-pay signals that British health
economics have done their best work. It began with Jones-Lee’s well-known studies on the valuation of human life [71] [see also 13, 30] but it soon moved on to specific issues of different medical treatments competing for the same funds. The three principal contenders for appraisal techniques have been cost-benefit analysis, cost-effectiveness analysis and cost-utility analysis, corresponding respectively to locative efficiency, technical efficiency and something like productive efficiency. Cost-benefit analysis is virtually excluded by the impossibility of pricing the consumption of health services in the NHS [72]. Cost-effectiveness analysis is perfectly feasible in principle but difficult to apply in practice because there is not hard evidence on the effectiveness of most medical treatments. Let us remember- and endlessly repeat to all those willing to listen - that fewer than one-quarter of all health care therapies now in use have been scientifically tested in double-blind field trials; the remaining two-thirds may well be efficacious but we do not know whether they do actually make the patient better [73, 74]; it is even conceivable that they make the patient worse [75]. We know that medical treatment for a given set of symptoms varies across countries, across regions within a country, and even across doctors in the same town in the same country [30, 76]. In short, to say that the health care sector is characterised by poor outcome data [73] it to labour the obvious. Medicine is a weak science – so is economics but medicine is probably the weaker of the two. Voltaire once defined a doctor as “a man who helps you to pass the time while Nature affects, the cure” despite all the progress made in medicine since the eighteenth century, we cannot entirely dismiss Voltaire’s cynical observation even now.

Mind you, if were provided with unambiguous indicators, of the effectiveness of various recommended cures for clearly diagnosed illnesses, it would still be true that effectiveness, as perceived by other is not the same thing as effectiveness perceived by ourselves; indeed, the effectiveness of the same medical care is bound to be perceived differently by different patient. What we look after, therefore, is what has been called cost-utility analysis according to which the individual patient passes judgment on what is an effective cure for whatever ails him or her [77]. The best-known example of cost-utility analysis is Alan’s William: QALY-measure or the extra “quality-adjusted-life-years”, that the patient can enjoy after treatment for a specific illness as evaluated by the patient him or herself [78]. A similar quality-quantity indicator is HYE or healthy-year-equivalents and DALY or disability-adjusted-life-years, with some arguing for the latter over QALYs on practical measurement grounds [79, 80]. In either case, what is involved is the increase in life expectancy from a particular treatment adjusted for the expected change in mobility, self-care, pain, discomfort, anxiety, depression, etc.

It would be true to say that until roughly 1980, it was either cost-benefit or cost-effectiveness analysis that dominated the set of techniques that British health economics employed to appraise alternative health care interventions. Drummond and others reviewed over 180 of these cost-benefit-cost effectiveness studies carried out before 1980 and five years later they updated these with a review of another 100 studies [81]. But after Williams [78] there was an overwhelming drift towards some kind of a look of cost-utility approach along the lines of QALY [77, 82]. It is clear that the cost QALY-criterion deliberately rejects the Pareto-principle of willingness and ability-to-pay for rationing health care and instead adopts a comparison of treatment-effects as perceived by patients themselves; since patients are competing with one another for a given budget, the ultimate,
The decision is left to the medical authorities and is made in terms of maximizing a set of QALYs per unit of medical costs.

The steady drift toward the approval of the QALY measures underlines my earlier contention that British health economics have travelled a long way from the Pareto doctrine and certainly further than the British economic profession as a whole. Some health economists have objected strenuously against the QALY – approach, not so much because it violates “welfarism” but because it neglects questions of equity [83-85]. But that appears to be a misunderstanding because QALYs can accommodate a wide variety of health dimensions and sources of valuation, and even the differential weighting of the benefits of treatment programmes according to who receives them. Their use implies no more than the concertalist position that the health of people, in the sense of both the length and the quality of life, is central to priority setting in health care and that, the bases for such priority-setting must be made as precise and explicit as possible [86]. On balance, and despite some criticisms in the early 1990s, there is little doubt that QALYs now rule the roost as the leading output measure for resource allocation in health care among British health economists [13, 30, 87, 89].

The normative welfare economics of health

Anyone thinking about the NHS is struck almost immediately by the fact that equality of treatment irrespective of ability-to-pay was a major source of inspiration for the creation of the system. It is hardly surprising, therefore, that British health economists have devoted a good deal of attention, and increasingly so in recent years, to the equity-efficiency – trade off implied by the NHS principle of rationing medical care by “need” to zero price.

The argument began in earnest with a paper by Mooney, Hall, Donaldson and Gerard on “Utilisation as measure of equity” [90]. This was answered by Culyer, van Doorslaer and Wagstaff [91], replied to by Mooney [92] and counter-replied by Culyer [93]. Mooney, Hall, Donaldson and Gerard denied that treating people with equal medical needs equally is the appropriate notion of equity in the deliverance of health care and argued instead that it should be the more modest objective of “equality of access” to medical facilities. Culyer et al attacked this argument because while equality of access is no doubt the distributional principle most congenial to Paretian welfare economics-let the patient be the judge of what is needed- the deliberate attempt of the NHS to sever the link between ability to pay and receipt of health care was equivalent to rejection of Paretian value judgements in the context of the health care. What is “just” or thought to be just is one thing and what accords more closely to Paretian principles is another “Equal treatment for equal medical need” or “horizontal equity” is, surely, what health authorities in the NHS are charged to carry out. In short, it is utilization, not access, that motivates British policy markers.

Despite all the to-ing and fro-ing of different definitions of equity [13, 30, 31, 34], this is largely a quarrel about words. The equality of access would imply a doctor and hospital with spare capacity within easy reach of every family in Britain and this simply is an unreachable ideal. Equal treatment for equal need would be an aim that we might conceivably attain if “need” were an objective concept that could be precisely defined in medical terms. But given the parlous state of medical knowledge, need is a subjective state of ill health that varies with the capacity...
of the health service to deal with sick patients [13, 88, 94]. In consequence, there is not much to choose between equal access to and equal use of medical services. The equity objective that the NHS actually achieves is “equal access for equal need” where need means the attainment of a standardised population mortality ratio for a region after consultations and sometimes protracted bargaining with local health authorities.

Normative arguments apart, there is the positive question whether the NHS actually succeeds in delivering equal treatment for equal need regardless of ability to pay. This is a subject that Julián Him Grand has virtually made his own [95]. He has shown, to simplify his findings, that the rich receive more health care than the poor. O’Donnell and Propper [96] have denied this but not convincingly so [97-99] and a fair r summary of the state of play in this area is that variations in the quality of treatment by income level are smaller in the U.K. than elsewhere but also that the NHS has never succeeded in providing health care when needed to anyone regardless of income, social class and occupation [34, 47]. Even the Ministry of Health [100] has admitted that the NHS has failed in close gap in longevity and other standard health indicators between high and low socio-economic groups.

Finally, and whatever the equality or inequality of treatment, there is the question of whether payment for health care via the tax system is biased against the poor or the rich. Putting together equity in delivery of health care and equity in finance, are the tax-financed systems of universal medical care in countries like the UK, Sweden and New Zealand more “progressive” than the systems of France, the Netherlands and Spain, and are all of these in turn more “progressive” than the predominantly privately financed systems of Switzerland and the United States? As one might have anticipated, the answer to that question is almost certainly, Yes [101-103]. In summary, British health economics have certainly not ignored the normative aspects of health care and, in some, sense this has been the area of their best contributions to the entire subject of health economics.

Reforming the NHS

Whatever the reasons for the creation of the NHS, one of its unintended consequences was to contain national health expenditure more effectively than any other private or social insurance system of health care finance, so much so that “spending too little” has become the battle-cry of every attack on the NHS [104] in the same way that “spending too much” is the battle-cry of every attack on the US system of user charges for health care. Is it really third-party payment problems that account for the run-away health care expenditures of the US or it is simply the result of the US demand for better-quality, technology-intensive medical care, in which case we ought to see a better record of health outcomes in the US than elsewhere? We have already touched on this question when we looked earlier at the evidence for supplier-induced demand but we ask it again in terms of international differences in health spending, on the one hand, and international differences in health outcomes on the other. There is a rich literature on this question by both American and British, health economics and, needless to say, it fails to reveal a simple, unambiguous conclusion except that the much higher level of American health expenditure fails to produce a corresponding improvement in the health status of the American population, if indeed any marked improvement over European levels. Of course, the difficulties of measuring a
population’s health status even by a large combination of indices allows considerable room for argument [105-109].

Nevertheless, the fact remains that the patterns of demand of health care in the US and in the UK are so strikingly different, no doubt because of the contrast in health care finance systems in the contrast in health care finance systems in the two countries, as to produce an ideological chasm that effectively - like to produce an ideological breach that indeed impedes fruitful discussion of advantages and disadvantages of the markets in health care. Alan Williams [110-112] has drawn a striking contrast between “egalitarianism” and “libertarianism” in terms of a set of social and political attitudes associated with alternative idealised and actual health care systems. This is a contrast which runs deep through all the debates concerning the recent introduction of “quasi markets” in the NHS.

Any economists brought up on the efficiency-equity trade-off realizes immediately that American medical cares is likely to show up well on efficiency grounds but poorly on equity grounds simply because it comes nearer than any other country to the principle of distributing medical care in accordance with purchasing power. Even so, the payee in the American system is the insurance company and the employer, not the patient, and the medical insurance as such, is as we know, a potent source of allocative inefficiency. We would expect American medical care to be highly equitable but potentially inefficient: a system in which there are not prices to register costly treatments is likely to show little relationship between medical inputs and output [47]. We are not surprised therefore, to learn of long waiting lists, despite empty beds - but we are surprised that evidence of inferior health care for poor people continues to appear.

Prior to, say the mid-1980s. British health economists generally exhibited sympathetic attitude at least to the broad aims of the NHS, although there was much concern about the inefficiencies generated by the system of rationing demand by waiting lists and the failure of the system to reveal the costs of alternative modes of treatment [32, 113]. Very few British health economists went so far as Institute of Economic Affairs authors in rejecting the NHS. root and branch. Indeed, over the years, say between the 1960s and 1980s, the drift of majority opinion moved if anything in favour of rather than against the NHS, possibly because the blatant inequities of free-market health care in the USA became increasingly obvious [114]. Then in 1985, Alan Enthoven published what proved to be an influential tract in which he recommended in “Internal Market Model” for the NHS according to which each District Health Authority (DHA) would receive a per capital revenue and capital allowance and would buy and sell services from other Districts and the private sector [115]. Four years later, the NHS white Paper Working for Patients [116] proposed the separation of “purchasers” and “providers” within the NHS, on essentially the lines outlined by Enthoven [117].

The introduction of quasi - markets in the NHS has produced a widespread -debate on the provision and finances of health care in Britain and a flurry of efforts to evaluate the impact of these reforms. Because the new
“purchasers” are both fund-holding GPs, DHAs and Family service Authorities and because the “providers” are both publics and private hospitals, not to mention ordinary GPs, dentists, opticians and pharmacists, the effects of the reforms are bound to be difficult to disentangle. Nothing like a consensus has yet been reached but a few general observations on the likely outcome of this debate are in order.

Firstly, we must underline the remarkable speed with which these rather drastic reforms have been introduced by the government without so much as a single field trial. When the US governments under presidents Nixon and Carter became alarmed about American health, care costs, they promoted the Rand Health Insurance Experiment, which assigned 2,000 non-elderly families from six different American states over a period of 3-5 year to a variety of medical insurance plans with various price tags, fee-for.-service insurance plans and group practice schemes, some of whom were randomly selected to serve as a control group [118]. Compared to this incidentally, the largest, longest-running social science research ever completed - we in Britain launched a serious overhaul of a forty-year-old system of health care provision and finance with little preparation, no research whatever, and pell-mell over a period of two years from stat to finish.

There has been some theoretical analysis applying principal-agency theory to the new contracts between providers and purchasers in the NHS [119, 120] and some effort to pin down precisely the information and incentives conditions that would be required for effective purchasing on the part of GPs and District Health Commissioners [121]. But the overwhelming opinion of all the health economists, who have looked carefully at these “quasi market” reforms is that at best they will encourage and evidence-based, cost-conscious health service that may improve outcomes for some patients registered with managerially competent GPs who can now obtain cheaper and better care from hospital, but at worst they will dramatically increase the cost of administration, produce no improvement in health care for the average patient, and may well destroy irremediably the morale of the NHS [33, 73, 122-129]. The reform certainly suffers from excessive optimism about the benefits of competition in health care. For example, there are two sets of “purchasers”: District General managers of fund-holding GPs, but since the two group have different priorities and since the lines of financial control over purchasing are unclear, it is very likely that the resulting coordination difficulties will never be decisively resolved. Also management decisions are now increasingly based on cost data and, since evidence on medical outcomes is ambiguous, doctors will be inevitably tempted to engage in “cream-skimming” implying a decline in over all health care [38]. As McGuire, Fenn and Mayhew [33] have said:

The reforms do not overcome the fundamental information asymmetries of that characterize the health care sector. And as such, it is not obvious that increased horizontal competition amongst providers will cut cost while simultaneously improving quality. Given the evidence from the USA, it seems likely that market initiatives may increase the costs of providing any given quality of care. It remains an open empirical question how costs and quality will react to the changed structure and incentives. The lack of precision that has accompanied the proposed reforms does not encourage the view that the new regulatory framework will be able to constrain the cost pressures that will be introduced with these reforms. Thus the impact of the reforms on the delivery of health care remains debatable.
Alan Blinder once laid down what he called Murphy Law of Economic Policy [130]: “Economists have the least influence on policy where they know the most; they have the most influence where they know the least and disagree most vehemently”. This Law certainly applies to health reform in NHS. All British health economics know that the special nature of the commodity called health care makes in difficult to impose a market in health care and even if artificially imposed, cannot be expected necessarily to lead to better outcomes than would a publicly provided, tax-financed NHS. In short, British health economics would have welcomed trial experiments ins certain localities with something like the US Health Maintenance Organizations for Britain, and even the creation of an internal market by the separation of providers and purchasers in a few DHAs, followed by a period of careful monitoring to check whether such a change would result in better outcomes. But instead the new reforms were instituted without experimental evidence and without so much as a sideways, glance at the opinions of British health economics thus confirming Murphy Law of Economic Policy: “Economists have the least influence on policy, where they know the most.”

One may ask: does it really matter whether the new reform will work? Surely, if they do not we can always roll them back in a few years? But such a reversal of policy may be impossible. It may well be that the NHS is riddled with shortcomings but in reforming it, we must be aware of “not letting the best be the enemy of the good”. If the first-best health care system we would all like to see ends up being a third-best inferior alternative to a second best pre-pre-reformed NHS the resulting situation may be irreversible. Health care systems are path dependent. History matters and history cannot be rolled back just because some ill-advised reforms proved, in the fullness of time, to be abortive, Let the health authorities of tomorrow take note.

**Conclusions**

So has there been progress in British health economics over the last quarter-century? Health economics may be said to have generated what the economics of education never succeeded in generating, namely, a paradigm-shift in the whole of economics. Out or health economics has come the economics of insurance, the economics of uncertainty. the market for “lemons”, missing markets, asymmetric information, etc. etc. The economics of education has remained since its inception in 1970, a somewhat remote outer region of economics; health economics, on the other hand, may be a topic familiar only to specialists but is theoretical concerns are the bread-and-butter of every economist.

Pareto-optimality, some British health economists have argued, is not applicable to health care. But what then of welfare economics? Are economists to be deprived of the judgement of “good” and “bad” in respect of economic events? This is a groundless fear because there is more to welfare economics than Pareto-optimality. There is the partial-equilibrium tradition of Marshall and Pigou, the applied welfare economics of cost-benefit analysis. What use any way is Paretian general-equilibrium welfare analysis except for purposes of generating purely formal. Invisible Hand theorems? Whenever we come to a practical problem, such as whether to introduce road pricing to deal with congestion problems, we inevitably resort to the Pigovian apparatus of private and social cost and benefits. The distinction between Paretian and Pigovian welfare economics is the same as that
between global and piecemeal “social engineering: what we want to know is not how to transform the entire world but how to make one corner of it a little better.

Paretian welfare economics is all about the causes of “market failure” Pigovian welfare economics is also about “government failure” and in a publicly provided, publicly financed health care system it is government failure that is the most relevant area of concern in welfare economics.

In appraising health care schemes like the recent NHS reforms, what we are left with are rough-and-ready, case-by-case qualitative judgements, which may be loose and imprecise but which have some chance of being right, And as Keynes once said: “it is better to be vaguely right than to be precisely wrong”. And it is vaguely right because the health economics of 1997 is immeasurably better informed both conceptually and empirically than was the health economics of 1970.

References


6. the concept of Pareto-optimality rests on there interrelated premises: (1) consumer sovereignty (2) no-paternalism-social welfare is the arithmetic sum of individual welfares: and (3) unanimity-only, unanimous reallocations of resources count as improvements in social welfare, that is, “thou shalt not make interpersonal comparisons of welfare”. The last one almost as troublesome for health economics as the first, ruling out as it does any externalities of consumption.

7. Paretian welfare economics is in any case condemned as irrelevant for other reasons. Since the postulate of unanimity would seem to rule out virtually all actual economics changes, the “new” welfare economics of Hicks-Kaldor fame supplemented Pareto by adding the notion of compensation payments according to which the potential gainers of a Pareto-optimos improvement could
bribe losers to accepted the changes while still being better off than before. However, to recommend such payment is to engage in interpersonal comparisons of utility because it amounts to equating the utilities of different individuals to money gains and losses. On the other hand, to leave the parties themselves to negotiate the appropriate level of compensation payments is to invite a strategic game of bluff and counter-bull which a game theory teaches us might well lead to stalemate. In short, if we rigorously uphold all three premises of Pareitian welfare economics, the set of feasible potential Pareto-improvements is either empty or vanishingly small—and this apart from the fact that the gainers and losers of most significant economics changes can only be identified at inordinate costs. Any applied welfare economics that deserves the name must make explicit interpersonal comparisons of welfare. This has been argued both by Williams, A. Cost-Benefit Analysis: Applied Welfare Economics or General Decision Aid? In: Williams, A., Gardina, E. (eds) *Efficiency in the Public Sector*. Edward Elgar Publishing: Aldershot, Hants: 1993: 65-79 and Culyer [88], with special reference to health economics.


19. *The new Palgrave a dictionary: of economics* (1987) does not contain and entry for “economics of education” although the *Economics of Education Review* like the *Journal of Health Economics* began publishing in 1981. Instead it refers the reader to an essay on “Human Capital” as if the economics of education were merely another name for human capital theory, a proposition which is roughly equivalent to equating labour economics with the marginal productivity theory of wages.


22. The only idea to have appeared with renewed vigour in the 1990s is the old one of income-contingent student loans in higher education, repaid (and this is the new twist) as a surcharge to national insurance contributions. See Barnes, J. and Barr, N. *Strategies for higher education: the alternative White Paper*. Aberdeen: The David Hume Institute, 1988; Barr, N. *Student Loans: The next step*. Aberdeen: The David Hume Institute, 1989: and West, E. G. *Britain’s Student Loan system: a world perspective: a critique*. London: Institute of Economic Affairs, 1994.


27. Weisbrod [51] conjectures that the explosion in health care expenditure in recent decades is the product of a retrospective cost-based finance system providing incentives to the R & D sector in...
contrast to the prospective payment system of primary and secondary education (essentially a fixed grant per child per school). If retrospective reimbursement had prevailed for schools, the private sector would have devoted more sources in improved educational diagnostic and learning techniques, in which case education like health care would have improved dramatically but so would educational expenditures. A corollary of this thesis is that the economics of education might have advanced like health economics because the cost explosion would have drawn new intellectual resources to education. (I owe this point to Diane Dawson.) But this assumes that progress in any sub-discipline of economics is a simple reflection of the political problems generated by the sector studied in that sub-discipline. This is too simplistic and almost anyone could name counter-examples. Don research programmes in economics have to internally generated momentum of their own?


48. Alan Williams has suggested to me that there may be supplier-reduced demand in the NHS because UK doctors are encouraged by their reward system to maximise inactivity and thus to suppress demand. Nevertheless medical, ethics encourages doctors to push the volume of medical care to the point of zero-marginal productivity and, in this way, there may be supplier-induced demand in the UK despite the lack of user charges.


52. Baumol, W.J. Children of performing arts, the economic dilemma: the climbing costs of health


65. As historian of the economic thought, I cannot forego the pleasure of commenting on an attempt to depict Adam Smith as a forerunner of health economics because of this alleged recognition of asymmetric information in any practical sense (to see West, E.G. Adam Smith. The man and his works. Indianapolis: Liberty Fund. 1976: and Mullahy, J. Adam Smith ace health economist. The oat-bran controversy was actually resolved in 1776. Journal of Health Economics 1990: 9: 367-69, on Smith’s praise of a wheat-rich diet against the standard medical opinion of his day.)


68. Feldstein, M. Economic analysis for Health Service


